Please cite as:
Enfield, N. J., & De Ruiter, J. P. (2003). The diff-task: A symmetrical dyadic multimodal interaction task. In N. J. Enfield (Ed.), *Field research manual 2003, part I: Multimodal interaction, space, event representation* (pp. 17-21). Nijmegen: Max Planck Institute for Psycholinguistics. doi:10.17617/2.877635.

REGULATIONS ON USE

Stephen C. Levinson and Asifa Majid

This website and the materials herewith supplied have been developed by members of the Language and Cognition Department of the Max Planck Institute for Psycholinguistics (formerly the Cognitive Anthropology Research Group). In a number of cases materials were designed in collaboration with staff from other MPI departments.

Proper citation and attribution

Any use of the materials should be acknowledged in publications, presentations and other public materials. Entries have been developed by different individuals. Please cite authors as indicated on the webpage and front page of the pdf entry. Use of associated stimuli should also be cited by acknowledging the field manual entry. Intellectual property rights are hereby asserted.

Creative Commons license

Background

The field manuals were originally intended as working documents for internal use only. They were supplemented by verbal instructions and additional guidelines in many cases. If you have questions about using the materials, or comments on the viability in various field situations, feel free to get in touch with the authors.

Contact

Email us via library@mpi.nl
Max Planck Institute for Psycholinguistics
P.O. Box 310, 6500 AH, Nijmegen, The Netherlands

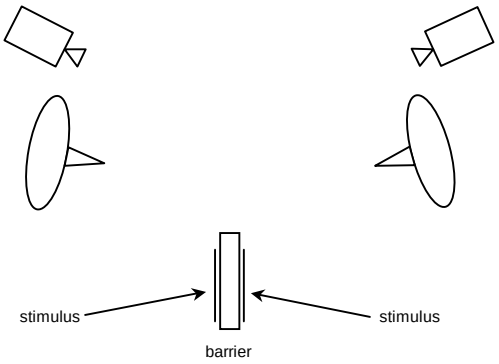
The Diff-Task: a symmetrical dyadic multimodal interaction task

N. J. Enfield, J. P. de Ruiter

Rationale: This task is a complement to the questionnaire 'Multimodal interaction in your field site: a preliminary investigation', above. The objective of the task is to obtain high quality video data on structured and symmetrical dyadic multimodal interaction. The features of interaction we are interested in include turn organization in speech and nonverbal behavior, eye-gaze behavior, use of composite signals (i.e. communicative units of speech-combined-with-gesture), and linguistic and other resources for 'navigating' interaction (e.g. words like *okay*, *now*, *well*, and *um*). The above questionnaire will provide much information on these issues, but due to the many differences in social organization and daily habits of people in different field sites, information from the questionnaire may not be directly comparable. It is known, for example, that behavior in interaction can be affected by the number of interactants, and by the presence or absence of potential foci of attention such as diagrammatic representations, pieces of writing, objects which figure in physical activities. Data from the questionnaire is more naturalistic than data elicited with this task, but provides less directly comparable data. The 'diff-task' will provide more directly comparable data by setting certain conditions in advance: the number of interactants (two), the number and type of external representations (see below), and the structure of the interaction (12 consecutive sub-tasks, each with defined completion points).

The diff-task is similar to traditional 'director-matcher' tasks, but is designed to avoid the interactional asymmetry inherent in the director-matcher design; i.e. the assignment of each individual to a unique and different communicative role. In the diff-task, participants have to solve a problem using referential communication, but each have precisely the same objective and precisely the same role.

Set-up: Ask two participants to get into a comfortable configuration for interaction. They can sit on the floor or on chairs, at a table, or whatever they like. They can be directly facing each other or at angles, whatever is comfortable for them just so long as their faces and torsos (including their 'gesture space') are in full view of the other. (This doesn't mean that they have to look at each other, but only that they should be able to look at each other if they choose to.) Try to make sure there is a clear space between the interactants (e.g. a table top or clear stretch of ground), where they may want to gesture or 'draw' (e.g. with their finger on a table top) as they talk. They can sit as close or as far apart as is comfortable, but don't forget to note down the details of how they are positioned. Then, a small barrier is set up at the middle distance point between them, but importantly it is placed to the side, and not directly in the path of sight from one to the other.

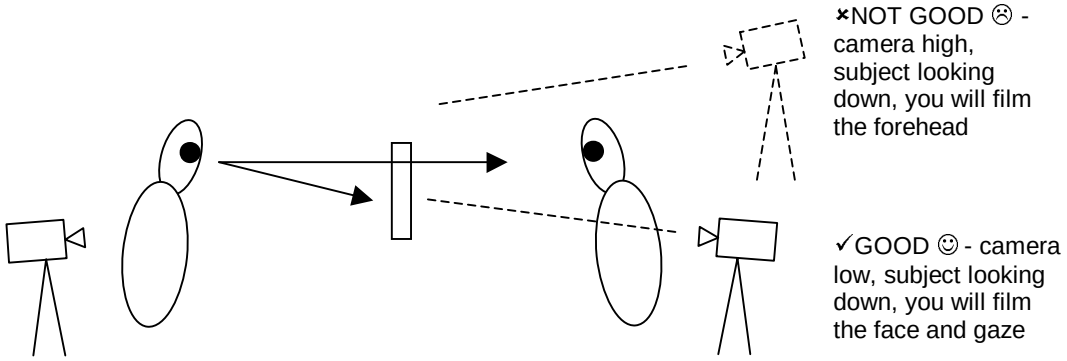


Overhead view of diff-task set-up: 'barrier' for the stimuli must *not* form a barrier to interactants' view of each other's faces and gestures.

On each side of the small barrier will be an illustration. (See below for instructions as to how to set this up.) Each participant will be able to see the illustration on their side of the barrier, but they cannot see the other's illustration (which is on the other side of the barrier). We recommend that you set up two video cameras, one for each interactant, with as good a view of their face and hands as possible. (The task can also be done with one camera; see below.) The shot needs to be fairly tight (i.e. need not be as wide as is customary for gesture studies, where one never knows how 'large' a speaker's gestures may be), and should include any possible surface between them which may be used for gesturing. Try to capture just one of the interactants (i.e. avoid capturing, say, the elbow of the second speaker in the corner of the shot). Here is how we composed the image in a pilot run of the task:

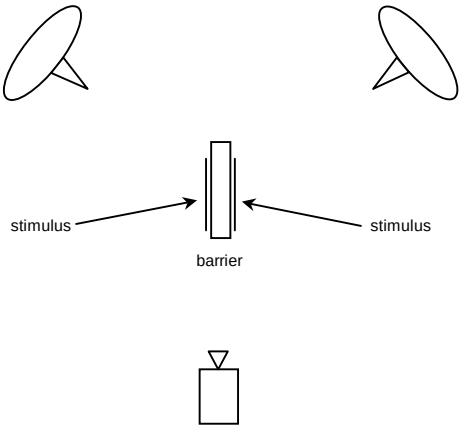
The technique of using two cameras requires that eventually the two recordings be precisely synchronized. The method to use is the 'clapboard': i.e. once both cameras are recording, bang two hard objects together so that from each camera view you will be able to see the two objects coming together, and discern the very frame in which the 'bang' sound happens. This should be done immediately after both cameras are rolling, and before you give the task instructions. DO NOT FORGET to do this before anything else. Also, you must bang the clapboard once again at the very end, so that we can check later if synchronization is accurate. Make sure it is visually captured by both cameras.

It is crucial that the camera be low enough that it sits approximately behind the line of gaze from the interactant to the illustration. If the speaker is looking down onto the illustration, and your camera is too high, then you will get a good view of their forehead, but you will not see what they are doing with their eyes. If your consultants are sitting on the floor and you can't get the camera low enough, then try raising the stimuli up high (e.g. using a chair), which would raise the line of the interlocutors' gaze.



Side view of task set-up: camera low so as to capture speaker's gaze.

Here is an alternative set-up if you only have one camera:



Alternative to two-camera set-up for the diff-task. The stimulus barrier splits the screen, and one participant can be clearly seen to each side. Both participants should be facing the stimulus at an angle, so that their faces will be visible to the camera.

The stimuli consist of 12 pairs of color illustrations, numbered 1a-12a and 1b-12b. They are photographs of simple arrays of objects. Half of them are relatively easy to name (cigarette, man, watch, rope, bottle, spectacles, knife) and half are relatively difficult to name (cylinder, letter O, microphone clip, little 'roof', wooden 'lemon', wooden 'banister', plastic cone). Here is an example pair of illustrations:

Pictures 1a and 1b. Spot the difference.

The sets of pictures are each provided in a clear plastic sleeve. The sleeve is attached to the barrier so that the first picture faces outwards. When the task starts, Speaker A sees Picture 1a, and Speaker B sees Picture 1b. The speakers are told that the pictures they see are different, and that their task is to find out how the pictures are different by talking to each other. The only real rule is that they are not allowed to look at each other's picture. They are finished with the first picture when they agree on what the difference is. They should then move on to the next pictures (i.e. 2a/b, and so on), by removing the outermost picture from the plastic sleeve, thus revealing the next one in order. Important: The participants should be asked *not* to report their finding to the experimenter after each time they have discovered a difference. You should avoid interaction with the participants during the task (except for clarification questions, etc.). We are interested in how the participants themselves 'navigate' the transitions from one sub-task to the next, and for this reason it is crucial that the fieldworker not do any of this navigation (e.g. saying 'okay, good, now let's move on to the next picture' or equivalent), unless absolutely necessary. You can make yourself unavailable while the task is under way by moving into the background, fiddling with equipment, etc.

Giving instructions to participants

The most important points to think of while preparing instructions in the target language for the diff-task are the following:

- The participants' task is to figure out ('discover', 'detect', 'assess') the way in which their own picture and the other person's picture are different, by talking to one another. (If you the experimenter are asked whether gesturing is allowed, say yes. But don't offer this information unless asked. People will gesture spontaneously.)
- Participants are not to look at each other's pictures. (Try to avoid having other people hanging around who will chime in.)
- The participants get no information from the experimenter about the *kind* of difference between the pictures. They have to figure that out for themselves.
- The participants should figure out the difference between all 12 pairs, self-paced, without communicating with the experimenter (i.e., you) about anything until the experiment is finished. Some of the pairs are easy to solve, some are quite difficult. If the participants are unable to figure out the difference after 5 minutes or so, give them clues. The task shouldn't go on for more than 30 minutes total. Note that it is not important to the success of the task that the speakers get all the answers. What's important is that they interact in trying.
- There is guaranteed to *be* a difference between the pictures (they are not being tricked). This may not need to be mentioned explicitly in the field, but experimentally savvy MPI subjects are known to have worried about this. In each case, there is only *one* difference, but this is not stated in the instructions. (Note that the position of the second hand on the watch in the stimulus pictures is indeed a difference, but this is never the difference we have incorporated. Consultants should, however, get a prize if they suggest this.)

There is no specific procedure after the experiment has been completed, but it will be interesting to see how people discuss the task, for example by looking over the pictures and checking the differences for themselves. You should continue to record this part of the interaction.

Here is an example instruction in English, from which you could base your target language instructions:

This is a game. You will each see a picture on this barrier, and your pictures will be different from each other's. The game is to find out how your picture and the picture of your partner are different. You cannot look at the other person's picture. The game is finished when you have done 12 pictures. So you start with the first picture, talk to each other and figure out how the pictures are different, and when you are sure you have agreed on how they are different, remove the picture you have just discussed, and go on to the next picture. You don't have to ask if it is okay to move on. Keep going until you have found the difference for all 12 picture pairs. When you have finished, please let me know.

DIFF TASK STIMULI

The content of the stimulus pictures is as follows:

Pair 01: Objects:: cylinder, cigarette, mike clip, man, watch, (in a, man to the left of watch, in b man to the right of watch): transformation: swap man & watch.

Pair 02: cigarette, banister, mike clip, rope, roofie, spectacles (seventies style spectacles in a, nineties style spectacles in b): transformation: change property of spectacles.

Pair 03: rope, cylinder, man, letter O (only present in a, not in b), cigarette, bottle, spectacles: transformation: remove letter O.

Pair 04: cylinder, rope, roofie, lemon (pointed side to the right in a, to the left in b), knife, spectacles: transformation: rotate lemon.

Pair 05: rope, man, cigarette, banister, spectacles: transformation: expand_outward, in a, objects are spaced much further apart than in b.

Pair 06: bottle, lemon, cone, watch , knife : in a: watch 'below' (before) knife, in a watch 'above' (behind) knife: transformation: swap watch & knife.

Pair 07: mike clip, letter O, banister, knife (blue handle in a, brown handle in b), transformation: change property of knife.

Pair 08: roofie, man, bottle, watch: In b, objects spaced close together, in a objects are much further apart., transformation: expand_across.

Pair 09: spectacles, letter O, mike clip, banister, cone, knife (blade points to the left in a, to the right in b): transformation: rotate knife 180 degrees.

Pair 10: lemon, knife, cylinder, roofie (roofie above blade of knife in a, above handle of knife in b): transformation: move roofie, .

Pair 11: man, cone, letter O, cigarette, cylinder, bottle (in a, bottle is closer to the viewer than cone and cigarette, in b it, is between cone and cigarette): transformation: move bottle.

Pair 12: watch, letter O, rope, cone, knife, bottle (a) cylinder (b): transformation: replace bottle by cylinder.